

1 1. A synthetic nucleic acid sequence which encodes a protein wherein at least one non-
2 common codon or less-common codon has been replaced by a common codon, and having one or
3 more of the following properties:

4 (i) the synthetic nucleic acid sequence comprises a continuous stretch of at least
5 90 codons all of which are common codons;

6 (ii) the synthetic nucleic acid sequence comprises a continuous stretch of common
7 codons, which continuous stretch includes at least 33% or more of the codons in the synthetic
8 nucleic acid sequence; or

9 (iii) wherein at least 94% or more of the codons in the sequence encoding the
10 protein are common codons and wherein the synthetic nucleic acid sequence encodes a protein of
11 at least about 90 amino acids in length.

1 2. The synthetic nucleic acid sequence of claim 1, wherein said synthetic nucleic acid
2 sequence encodes a protein wherein at least ~~one~~ non-common codon or less-common codon has
3 been replaced by a common codon, and ~~wherein~~ the synthetic nucleic acid sequence comprises a
4 continuous stretch of at least 90 codons all of which are common codons.

1 3. The synthetic nucleic acid sequence of claim 1, wherein said synthetic nucleic acid
2 sequence encodes a protein wherein at least one non-common codon or less-common codon has
3 been replaced by a common codon, and wherein ~~the~~ synthetic nucleic acid sequence comprises a
4 continuous stretch of common codons, which continuous stretch includes at least 33% or more of
5 the codons in the synthetic nucleic acid sequence.

1 4. The synthetic nucleic acid sequence of claim 1, wherein said synthetic nucleic acid
2 sequence encodes a protein wherein at least one non-common codon or less-common codon has
3 been replaced by a common codon, and wherein at least 94% or more of the codons in the
4 sequence encoding the protein are common codons and wherein the synthetic nucleic acid
5 sequence encodes a protein of at least about 90 amino acids in length.

1 5. The nucleic acid sequence of claim 1, wherein the continuous stretch occurs in a
2 nucleic acid sequence which is selected from a group of sequences consisting of a sequence of a
3 pre-pro-protein; a sequence of a pro-protein; a sequence of a mature protein; a "pre" sequence of
4 a pre-pro-protein; a "pre-pro" sequence of a pre-pro-protein; a "pro" sequence of a pre-pro or a
5 pro-protein; or a portion of any of the aforementioned sequences.

1 6. The nucleic acid sequence of claim 1, wherein the continuous stretch comprises at
2 least 95 common codons.

1 7. The nucleic acid sequence of claim 1, wherein the nucleic acid comprises at least 30
2 non-common or less-common codons, these codons having been replaced with common codons.

1 8. The nucleic acid of claim 1, wherein the number of non-common or less-common
2 codons replaced or remaining is less than 15.

1 9. The nucleic acid of claim 1, wherein the non-common and less-common codons, taken
2 together, replaced or remaining, are equal or less than 6% of the codons in the synthetic nucleic
3 acid sequence.

1 10. The nucleic acid of claim 1, wherein all of the non-common or less-common codons
2 of the synthetic nucleic acid sequence encoding a protein have been replaced with common
3 codons.

1 11. The nucleic acid of claim 1, wherein all of the non-common and less-common
2 codons of the synthetic nucleic acid sequence encoding a protein have been replaced with
3 common codons.

1 12. The nucleic acid of claim 1, wherein the nucleic acid sequence encodes a protein of
2 at least about 105 amino acids in length.

1 13. The nucleic acid of claim 1, wherein at least 96% of the codons in the synthetic
2 nucleic acid sequence are common codons.

1 14. The nucleic acid of claim 1, wherein at least 98% of the codons in the synthetic
2 nucleic acid sequence are common codons.

1 15. A synthetic nucleic acid sequence which encodes Factor VIII, wherein at least one
2 non-common codon or less-common codon has been replaced by a common codon and wherein
3 the synthetic nucleic acid has one or more of the following properties: it has a continuous stretch
4 of at least 90 codons all of which are common codons; it has a continuous stretch of common
5 codons which comprise at least 33% of the codons of the synthetic nucleic acid sequence; at least
6 94% or more of the codons in the sequence encoding the protein are common codons and the
7 synthetic nucleic acid sequence encodes a protein of at least about 90 amino acids in length; it is
8 at least 80 base pairs in length.

1 16. The synthetic nucleic acid sequence of claim 15 where the factor VIII protein has
2 one or more of the following characteristics:

3 a) the B domain is deleted (BDD factor VIII);

4 b) it has a recognition site for an intracellular protease of the PACE/furin class;

5 or

6 c) it is inserted into a non-transformed cell.

1 17. The synthetic nucleic acid sequence of claim 15, wherein the number of non-
2 common or less- common codons replaced or remaining is less than 15.

1 18. The synthetic nucleic acid sequence of claim 15, wherein the number of non-
2 common or less- common codons replaced or remaining, taken together, are equal or less than
3 6% of the codons in the synthetic nucleic acid sequence.

1 19. The synthetic nucleic acid sequence of claim 15, wherein all non- common or less-
2 common codons are replaced with common codons.

1 20. The synthetic nucleic acid sequence of claim 15, wherein all non- common and less-
2 common codons are replaced with common codons.

1 21. The synthetic nucleic acid sequence of claim 15, wherein at least 96% of the codons
2 in the synthetic nucleic acid sequence are common codons.

1 22. The synthetic nucleic acid sequence of claim 15, wherein at least 98% of the codons
2 in the synthetic nucleic acid sequence are common codons.

1 23. The synthetic nucleic acid sequence of claim 15, wherein all of the codons are
2 replaced with common codons.

1 24. A synthetic nucleic acid sequence which encodes Factor IX, wherein at least one
2 non-common codon or less-common codon has been replaced by a common codon and wherein
3 the synthetic nucleic acid has one or more of the following properties: it has a continuous stretch
4 of at least 90 codons all of which are common codons; it has a continuous stretch of common
5 codons which comprise at least 33% of the codons of the synthetic nucleic acid sequence; at least
6 94% or more of the codons in the sequence encoding the protein are common codons and the
7 synthetic nucleic acid sequence encodes a protein of at least about 90 amino acids in length; it is
8 at least 80 base pairs in length.

1 25. The synthetic nucleic acid sequence of claim 24, wherein the factor IX protein has
2 one or more of the following characteristics:

3 a) it has a PACE/furin site at a pro-peptide mature protein junction; or

4 b) is inserted into a non-transformed cell.

1 26. The synthetic nucleic acid sequence of claim 24, wherein the number of non-
2 common or less- common codons replaced or remaining is less than 15.

1 27. The synthetic nucleic acid sequence of claim 24, wherein the number of non-
2 common or less- common codons replaced or remaining, taken together, are equal or less than
3 6% of the codons in the synthetic nucleic acid sequence.

1 28. The synthetic nucleic acid sequence of claim 24, wherein all non- common or less-
2 common codons are replaced with common codons.

1 29. The synthetic nucleic acid sequence of claim 24, wherein all non- common and less-
2 common codons are replaced with common codons.

1 30. The synthetic nucleic acid sequence of claim 24, wherein at least 96% of the codons
2 in the synthetic nucleic acid sequence are common codons.

1 31. The synthetic nucleic acid sequence of claim 24, wherein at least 98% of the codons
2 in the synthetic nucleic acid sequence are common codons.

1 32. The synthetic nucleic acid sequence of claim 24, wherein all of the codons are
2 replaced with common codons.

1 33. A vector comprising the synthetic nucleic acid sequence of claim 1, 15, or 24.

1 34. A cell comprising the nucleic acid sequence of claim 1, 15, or 24.

1 35. A method for preparing a synthetic nucleic acid sequence which is at least 90 codons
2 in length, comprising:

3 identifying a non-common codon and a less-common codon in a non-optimized
4 gene sequence which encodes a protein; and

5 replacing at least 94% of the non-common and less-common codons with a
6 common codon encoding the same amino acid as the replaced codon.

1 36. The method of claim 35, wherein at least 96% of the non-common and less-common
2 codons are replaced with a common codon encoding the same amino acid as the replaced codon.

1 37. The method of claim 35, wherein at least 98% of the non-common and less-common
2 codons are replaced with a common codon encoding the same amino acid as the replaced codon

1 38. The method of claim 35, wherein the nucleic acid sequence encodes a protein of at
2 least about 105 or more codons in length.

1 39. A method for making a nucleic acid sequence which directs the synthesis of an
2 optimized message of a protein of at least 90 amino acids comprising:

3 synthesizing at least two fragments of the nucleic acid sequence, wherein the two
4 fragments encode adjoining portions of the protein and wherein both subunits are mRNA
5 optimized; and

6 joining the two fragments such that a non-common codon is not created at a
7 junction point, thereby making the mRNA optimized nucleic acid sequence.

1 40. The method of claim 39, wherein the two fragments are joined together such that a
2 unique restriction endonuclease site is not created at the junction point.

1 41. The method of claim 39, wherein the two fragments are joined together such that a
2 unique restriction site is created.

1 42. The method of claim 39, wherein three fragments of the nucleic acid sequence are
2 synthesized.

1 43. The method of claim 39, wherein the synthetic nucleic acid sequence encodes a
2 protein of 105 or more codons in length.

1 44. The method of claim 39, wherein 96% of the codons in the synthetic nucleic acid
2 sequence are common codons.

1 45. The method of claim 39, wherein 98% of the codons in the synthetic nucleic acid
2 sequence are common codons.

1 46. The method of claim 39, wherein all of the codons in the synthetic nucleic acid
2 sequence are common codons.

1 47. The method of claim 39, wherein the number of codons which are not common
2 codons is equal to or less than 15.

1 48. The method of claim 39, wherein each fragment is at least 30 codons in length.

1 49. A method of providing a subject with a protein or polypeptide, comprising:
2 providing a synthetic nucleic acid sequence that can direct the synthesis of an
3 optimized message for a protein or polypeptide;
4 introducing the synthetic nucleic acid sequence into the subject; and
5 allowing the subject to express the protein or polypeptide, thereby providing the
6 subject with the protein.

1 50. The method of claim 49, wherein the synthetic nucleic acid is introduced into a cell.

1 51. The method of claim 50, wherein the cell can be an autologous, allogenic, or
2 xenogeneic cell.

1 52. The method of claim 50 wherein the cell is a fibroblast, a hematopoietic stem cell, a
2 myoblast, a keratinocyte, an epithelial cell, an endothelial cell, a glial cell, a neural cell, a cell
3 comprising a formed element of the blood, a muscle cell and precursors of these somatic cells.

1 53. The method of claim 49, wherein the codon optimized synthetic nucleic acid
2 sequence can be inserted into the cell *ex vivo* or *in vivo*.

1 54. The method of claim 49, wherein at least 94%, or all of the codons in the synthetic
2 nucleic acid sequence are common codons.

1 55. The method of claim 49, wherein at least 96%, or all of the codons in the synthetic
2 nucleic acid sequence are common codons.

1 56. The method of claim 49, wherein at least 98%, or all of the codons in the synthetic
2 nucleic acid sequence are common codons.

1 57. The method of claim 49, wherein the number of codons which are not common
2 codons is equal to or less than 15.

1

1 58. A method for preparing a synthetic nucleic acid sequence encoding a protein which
2 is at least 90 codons in length, comprising identifying non-common codon and less-common
3 codons in the non-optimized gene encoding the protein and replacing at least 94% or more of the
4 non-common and less-common codons with a common codon encoding the same amino acid as
5 the replaced codon.

1 59. A primary or secondary cell of vertebrate origin having an exogenous synthetic
2 nucleic acid sequence which encodes a protein or a polypeptide wherein at least one non-
3 common codon or less-common codon has been replaced by a common codon and wherein the
4 synthetic nucleic acid has one or more of the following properties: it has a continuous stretch of
5 at least 90 codons all of which are common codons; it has a continuous stretch of common
6 codons which comprise at least 33% of the codons of the synthetic nucleic acid sequence; at least
7 94% or more of the codons in the sequence encoding the protein are common codons and the
8 synthetic nucleic acid sequence encodes a protein of at least about 90 amino acids in length; it is
9 at least 80 base pairs in length and which is free of unique restriction endonuclease sites that
10 would occur in the message optimized sequence; and

11 DNA sequences, sufficient for expression of the exogenous synthetic DNA in the
12 transfected primary or secondary cell;

13 the primary or secondary cell capable of expressing the protein or polypeptide product.

1 60. The primary or secondary cell of claim 59, wherein the exogenous synthetic nucleic
2 acid is transfected into the cell.

1 61. The primary or secondary cell of claim 59, wherein the exogenous synthetic nucleic
2 acid sequence is stably integrated into its genome.

1 62. The primary or secondary cell of claim 59, wherein the exogenous synthetic nucleic
2 acid is present in the cell in an episome.

63. The primary or secondary cell of claim 59, wherein the DNA sequence sufficient for
expression of the exogenous synthetic nucleic acid is of non-viral origin.